Understanding latent sector errors and how to protect against them

Bianca Schroeder, Sotirios Damouras, Phillipa Gill

UNIVERSITY OF TORONTO

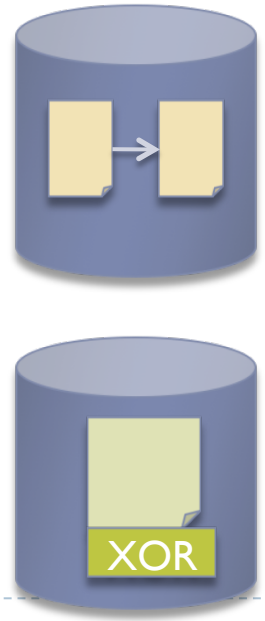
Motivation

- What is a latent sector error (LSE)?
  - Individual sectors on a drive become inaccessable (media error)
- Prevalence?
  - 3.5% of drives experience LSE(s) [Bairavasundaram2007]
    - 7-9% for some disk models!
- Consequence of an LSE?
  - In a system without redundancy: **data loss**
  - In RAID-5, if discovered during reconstruction: **data loss**
    - One of the main motivations for RAID-6
    - Growing concern with growing disk capacities

How to protect against them?

- Periodic scrubbing
  - Proactively detect LSEs and correct them.

- Intra-disk redundancy
  - Replicate selected metadata [e.g. FFS]
  - Add parity block per file [e.g. Iron file systems]
  - Add parity block per group of sectors [Dholak.08]

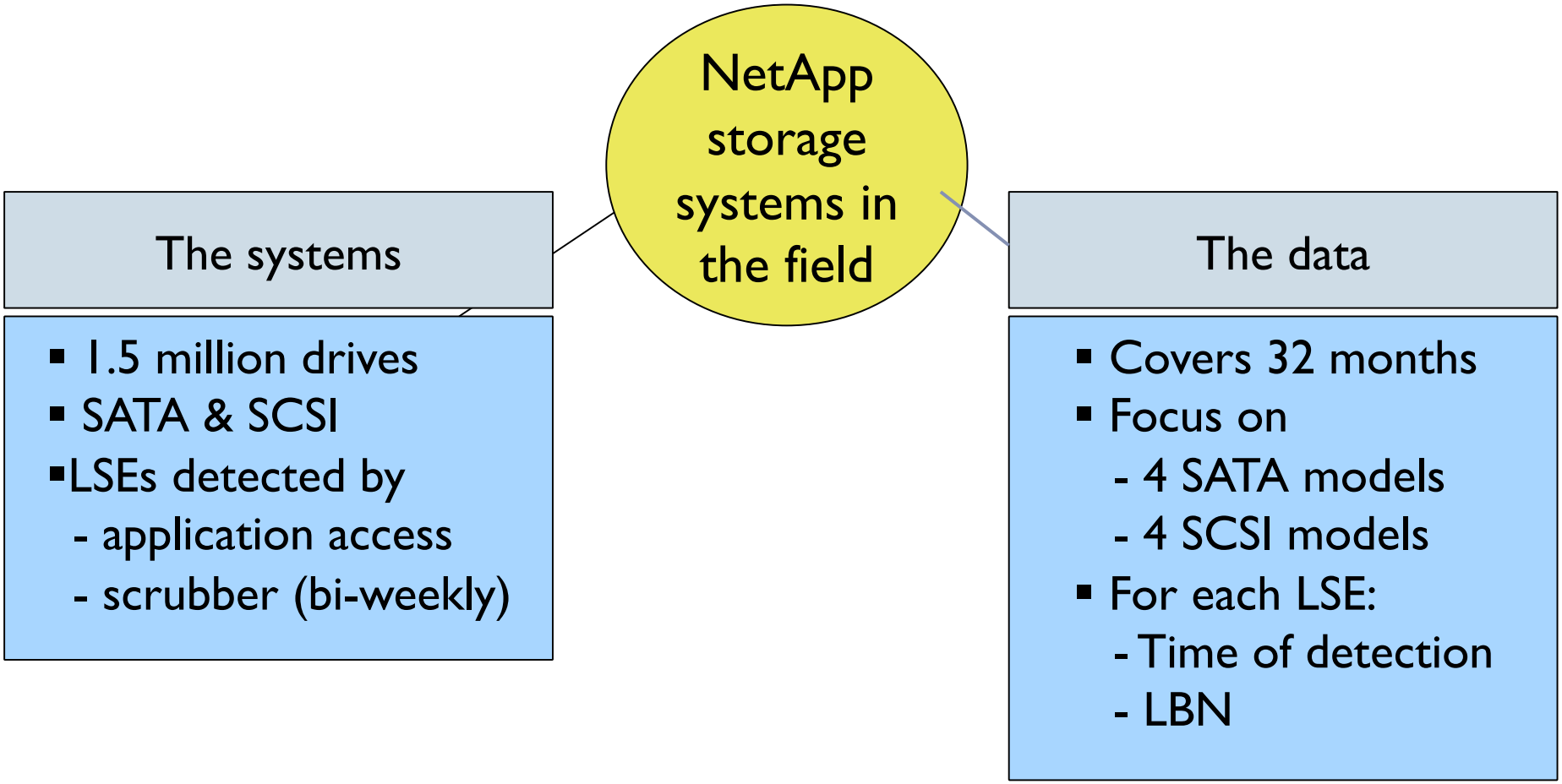


Our goal

- Understand potential of different protection schemes
- Understand characteristics of LSEs
  - From point of view of protection

- How?
  - Using real data from production machines
  - Subset of data in Bairavasundaram et al. (Sigmetrics'07)
  - **Thanks for sharing!**

The data


NetApp storage systems in the field
The systems
1.5 million drives
SATA & SCSI
LSEs detected by
- application access
- scrubber (bi-weekly)
The data
Covers 32 months
Focus on
- 4 SATA models
- 4 SCSI models
For each LSE:
- Time of detection
- LBN

How effective are protection schemes?

- Scrubbing
- Intra-disk redundancy

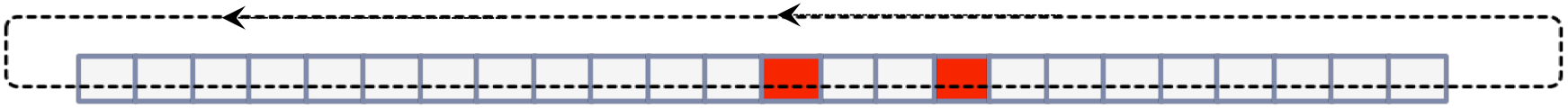
Scrubbing

- Why?

- Detect and correct errors early
- Reduces probability to encounter LSE during RAID reconstruction

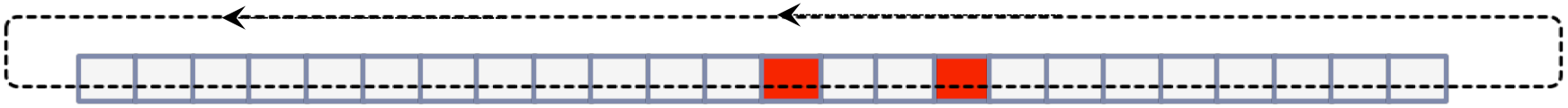
Scrubbing

- Standard sequential scrubbing

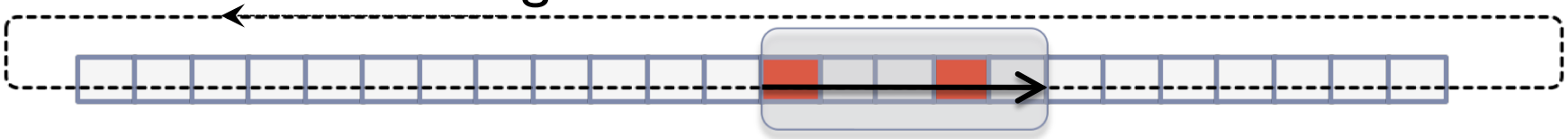


Scrubbing

- Standard sequential scrubbing

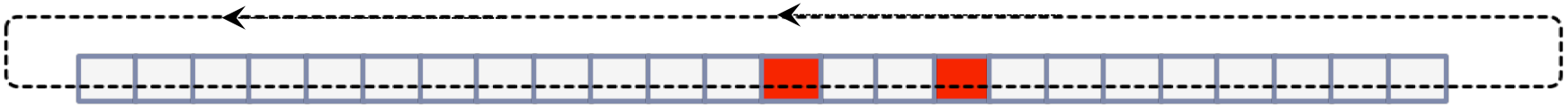


- Localized scrubbing

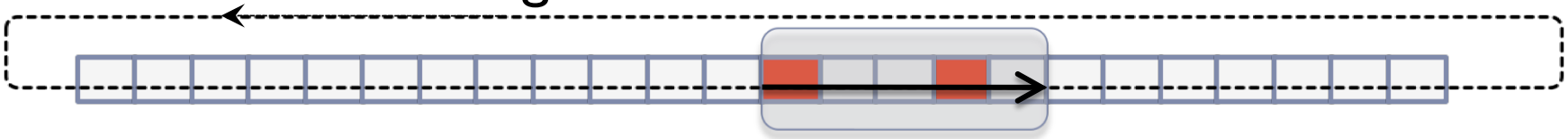


Scrubbing

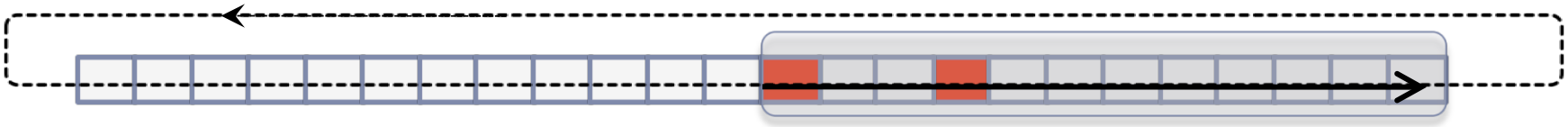
- Standard sequential scrubbing



- Localized scrubbing

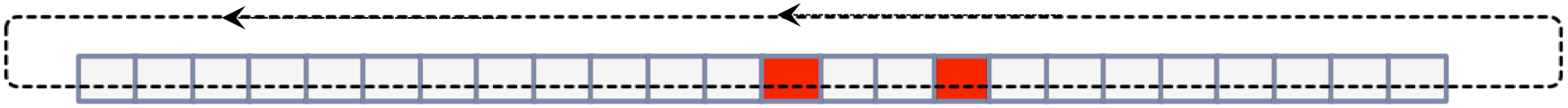


- Accelerated scrubbing

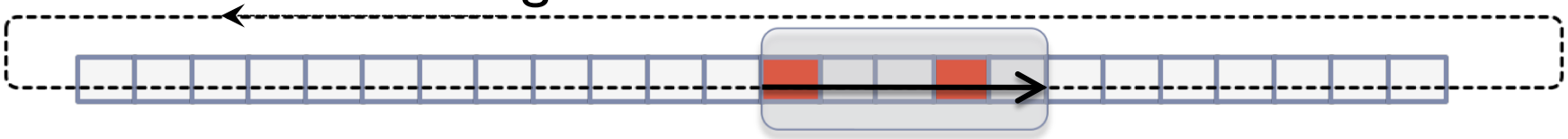


Scrubbing

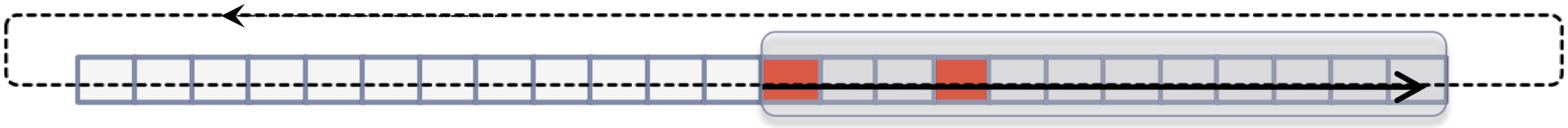
- Standard sequential scrubbing



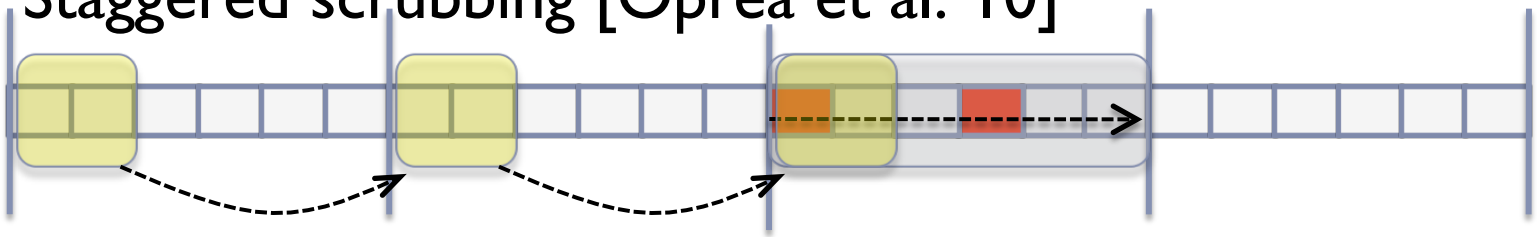
- Localized scrubbing



- Accelerated scrubbing



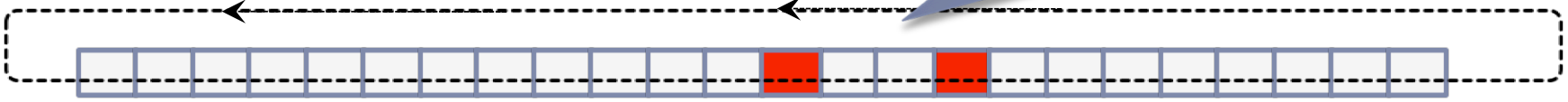
- Staggered scrubbing [Oprea et al.'10]



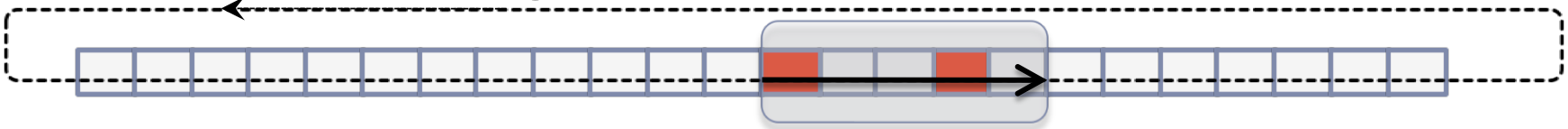
Scrubbing

How do those approaches perform in practice, i.e. on real-world data?

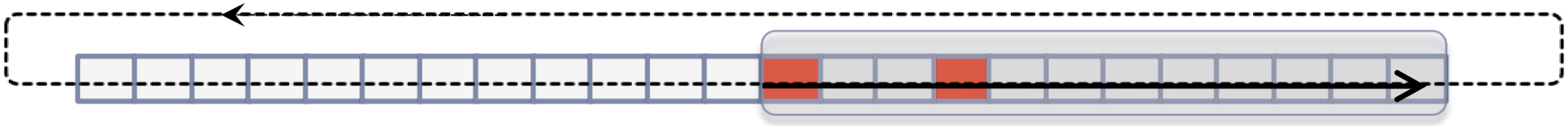
- Standard sequential scrubbing



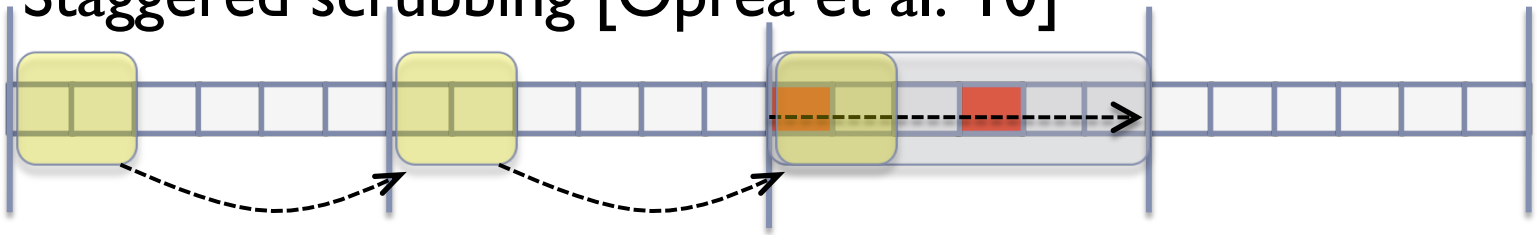
- Localized scrubbing



- Accelerated scrubbing

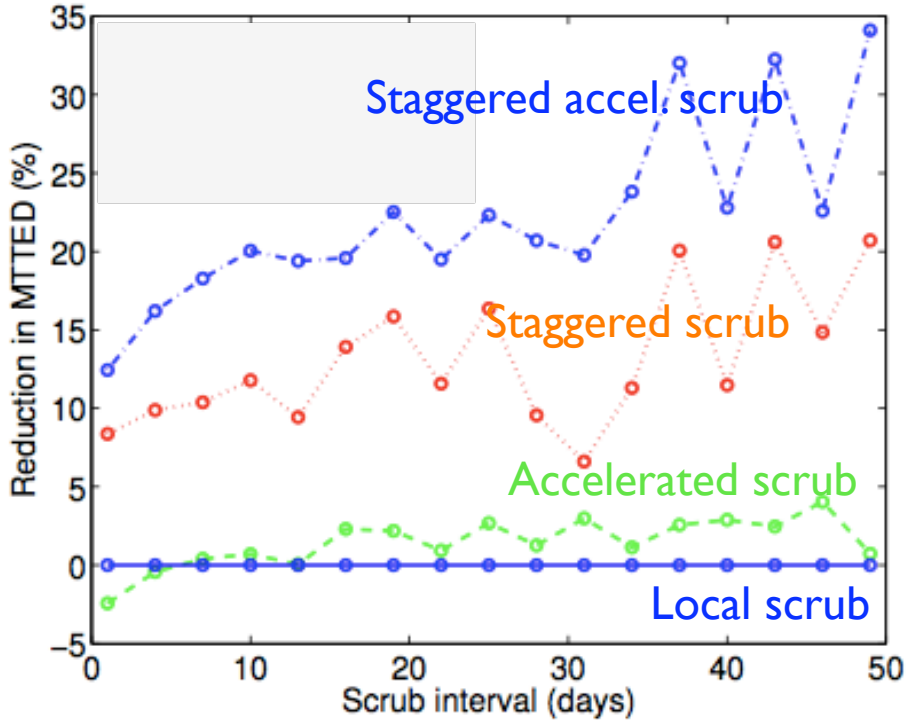


- Staggered scrubbing [Oprea et al.'10]



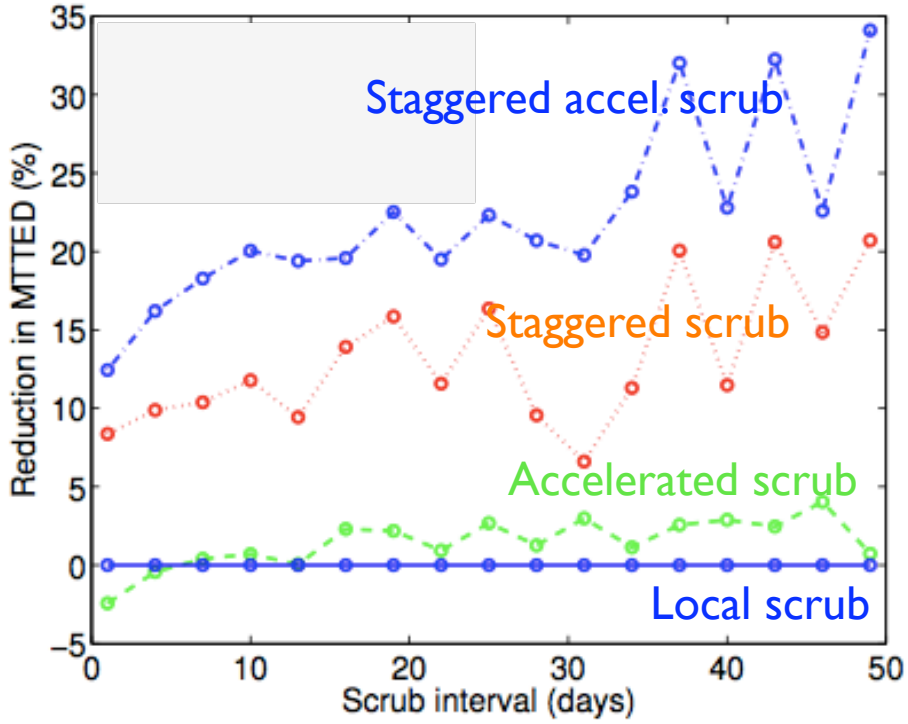
- Accelerated staggered scrubbing

Scrubbing: Evaluation on NetApp data



- No significant improvement from local & accelerated scrubs
  - They don't reduce the time to detect whether there are any errors
  - Errors are close in space, so even standard scrub finds them soon

Scrubbing: Evaluation on NetApp data



- 10-35% improvement with staggered scrubs!
  - Even better than the original paper claims!
  - Without introducing any overheads or additional reads
  - Relatively insensitive to choice of parameters

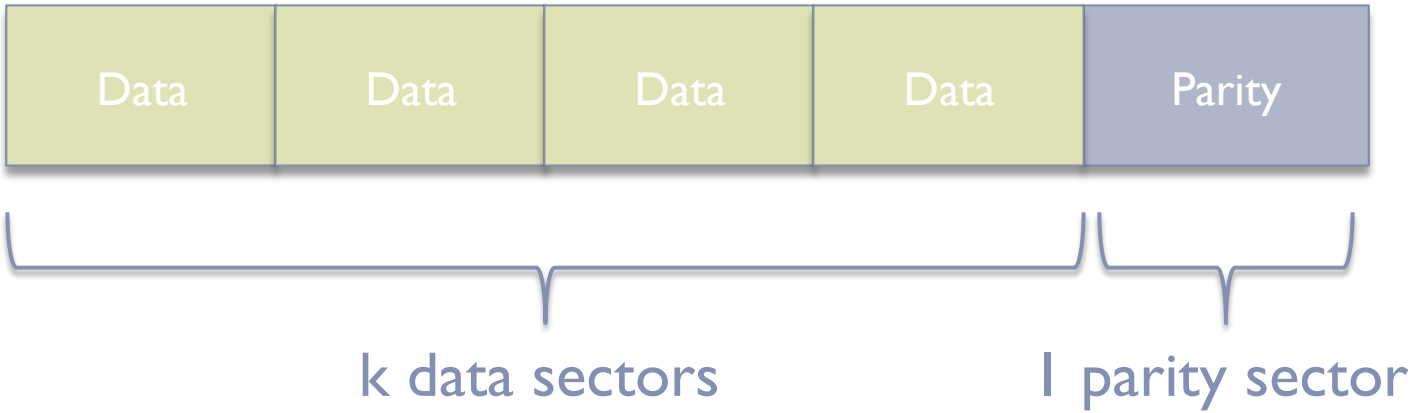
Intra-disk redundancy

- Why?
  - Recover LSEs in systems without redundancy
  - Recover LSEs during reconstruction in RAID-5

- Goal:
  - Evaluate potential protection
    - What fraction of errors could be recovered
  - Qualitative discussion of overheads

Intra-disk redundancy

- Simplest scheme: Single Parity Check (SPC)
- Can recover up to one LSE per parity group



- Results from evaluation on Netapp data:
  - 25-50% of drives have errors that SPC cannot recover

- Consider stronger schemes?

Stronger schemes?

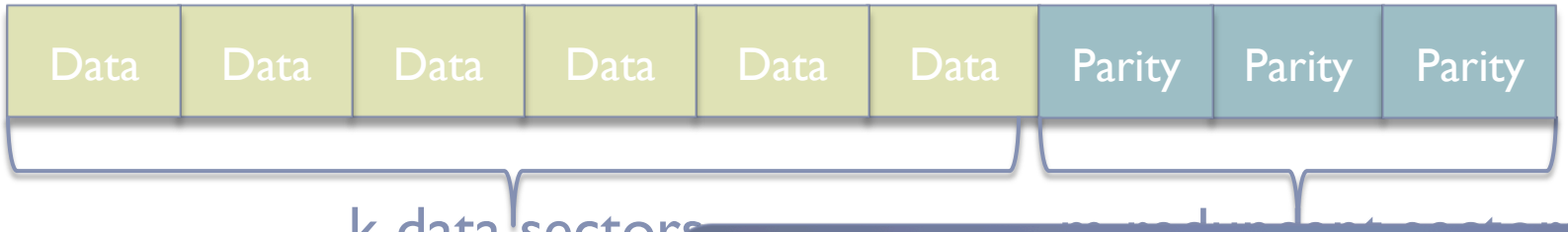
- Additional parity => additional overhead in updating parity
- When would that be interesting?

- In environments
  - … like archival systems, that don't have updates and don't like scrubs since they require powering up the system
  - … with read-mostly workloads, i.e. parity updates are rare
  - … for selected critical data on a drive, such as meta-data

Inter-leaved Parity Check (IPC) [Dholakia08]

- Requires only 1 parity update per data update
- Can tolerate up to m consecutive errors

Inter-leaved Parity Check (IPC) [Dholakia08]



- Claim: Achieves protectio
  - MDS=Maximum distance s
  - Expensive, but can tolerate

- Results differ from [Dholakia08]
- Importance of real-world data.
- Paper provides models & parameters

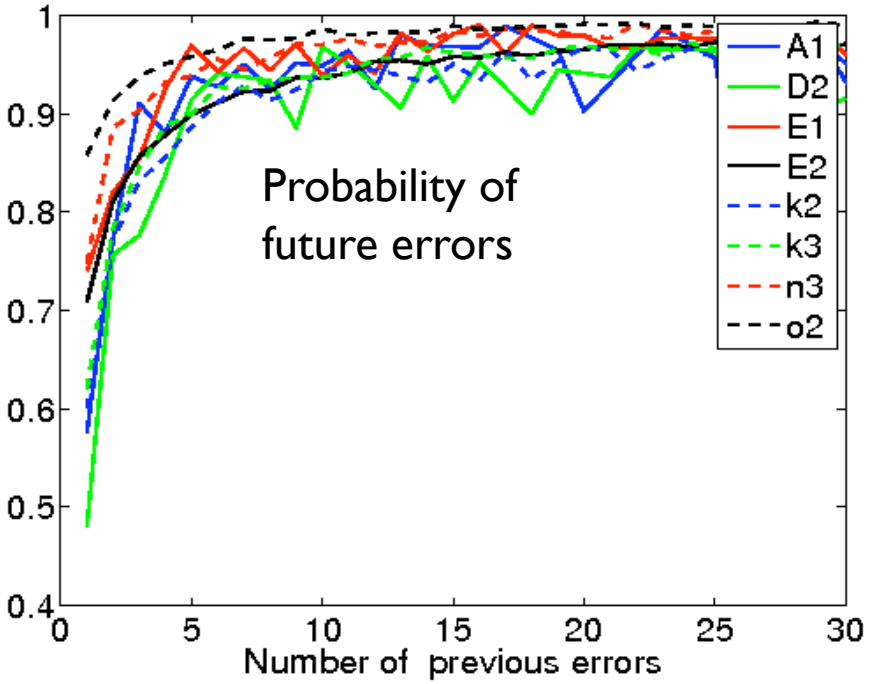
- Results: (from evaluation on Ne p data)
  - Far weaker than MDS!
  - Not significantly better than SPC

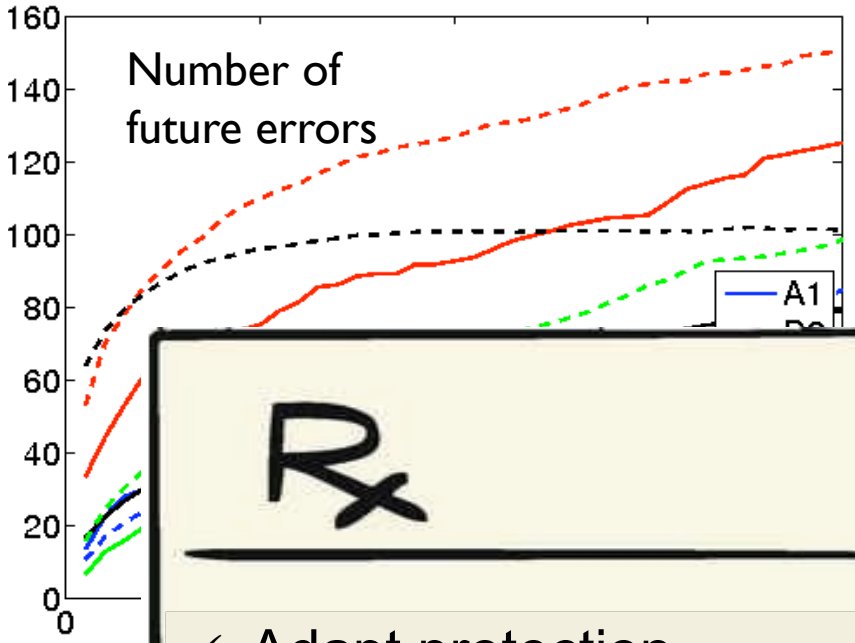
- Implications
  - Need different ideas for improving on SPC
  - Maybe reuse ideas from RAID-6? (see paper for details & results)

Questions unanswered …

- What level of protection to use when?
  - E.g. what is the right scrub frequency?
  - Depends on error probability at a given time

Do previous errors predict future?



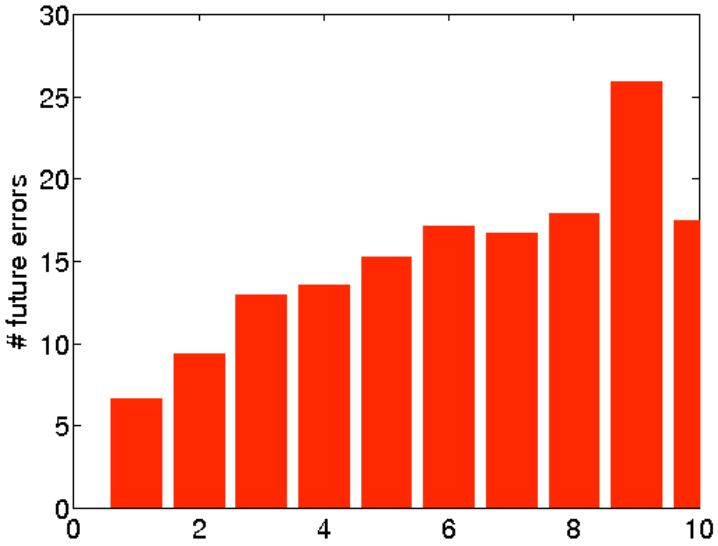


- Many previous errors
  - => higher chance of future error
  - => higher number of future erro
- Big differences between models

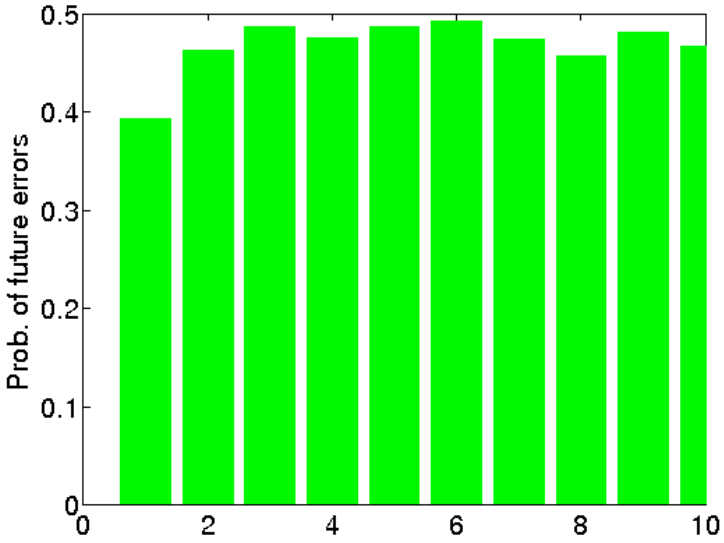
Rx

- ✓ Adapt protection based on previous errors
- ✓ Know your patient ..

Does *first* error interval predict future?



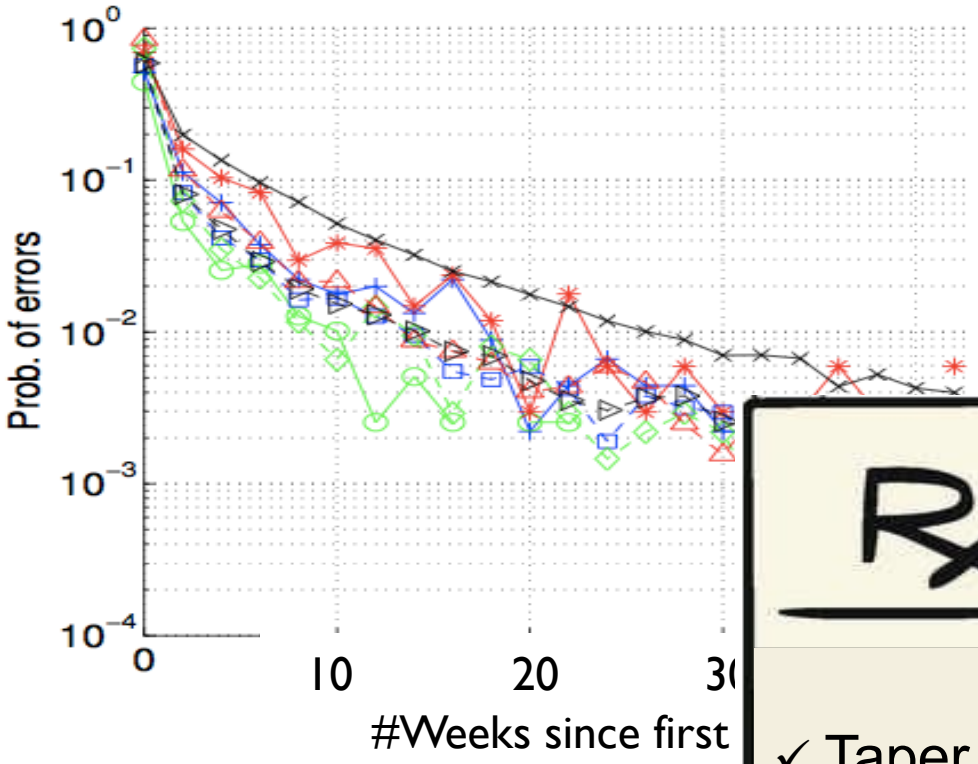
#errors in first scrub with errors



#errors in first scrub with errors

- Number of errors in first error interval:
  - Do increase expected number of future errors
  - Don't significantly increase probability of future occurrence

For how long are probabilities increased?



- Exponential drop-off, but still signifi
- Independent of number of errors i

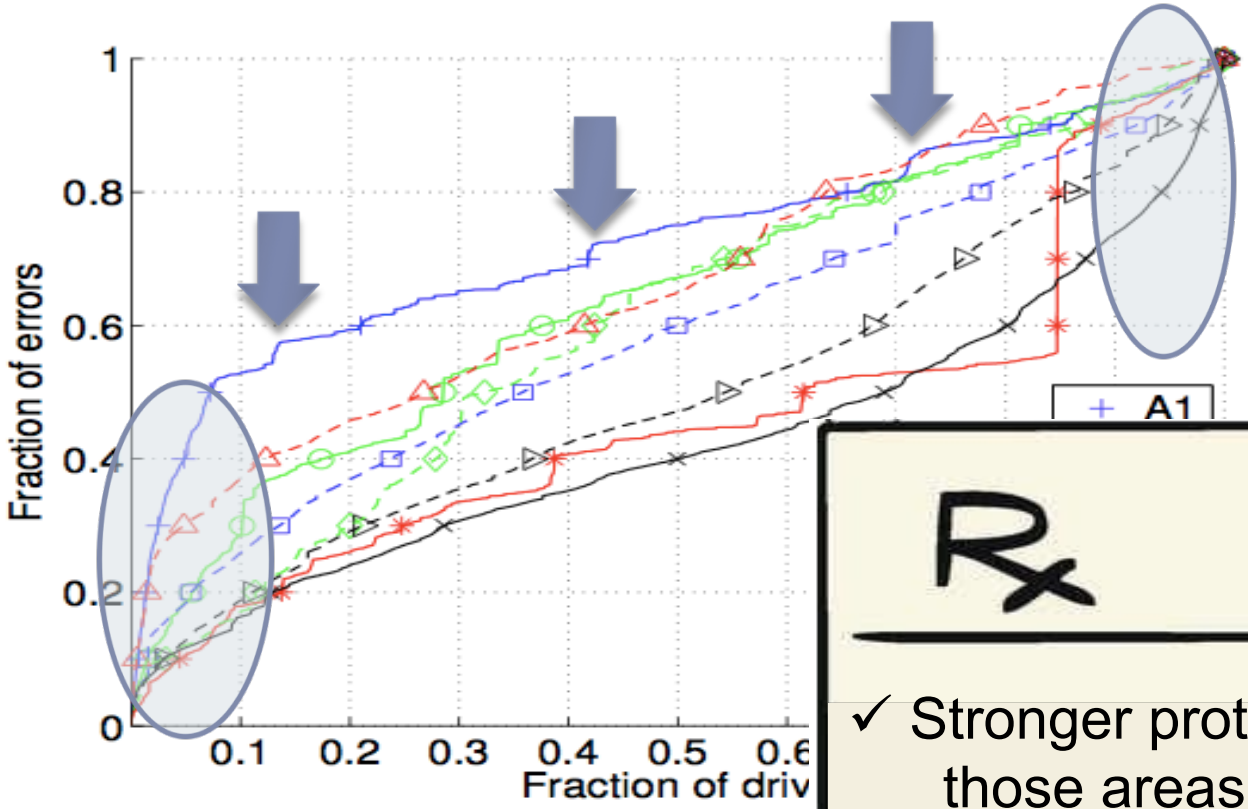
Rx

✓ Taper off added protection over time, e.g. reduce scrub rate

Questions unanswered …

- What level of protection to use when?
  - What is the error probability at a given time?
- What level of protection to use where?
  - Are all areas of the drive equally likely to develop errors?

Where on the drive are errors located?



- Up to 50% of errors concentrated in t
- Also increased probability in some oth

Rx

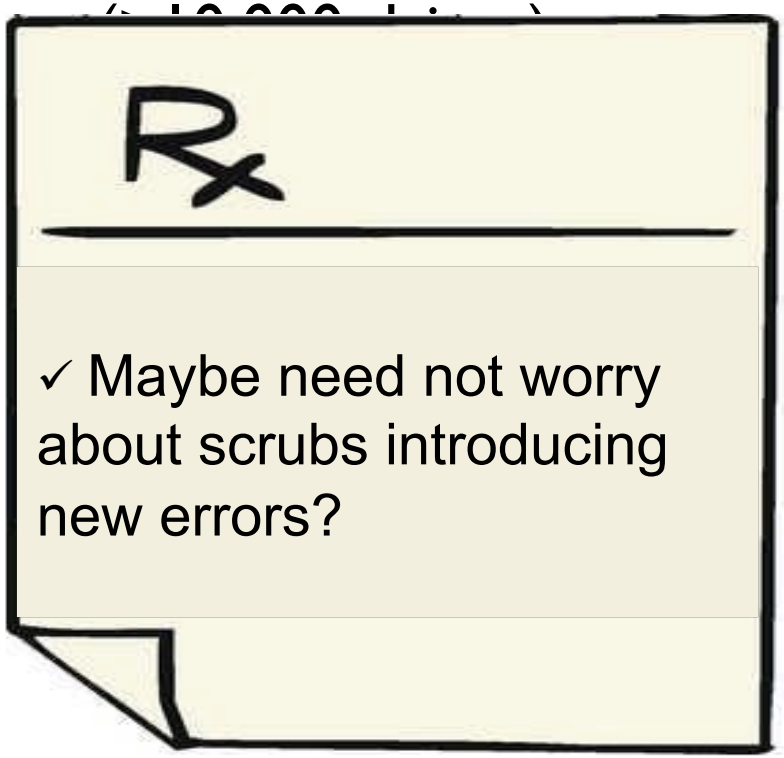
- ✓ Stronger protection for those areas
- ✓ Don't use for important data

Questions unanswered …

- What level of protection to use when?
  - What is the error probability at a given time?
- Same protection scheme across entire drive?
  - Are all parts equally likely to develop errors?
- Scrubbing potentially harmful?
  - Do additional read operations increase error rate?

Does utilization affect LSEs?

- Collected data in Google data cen
  - Number of LSEs
  - Number of reads & number of
- Results:
  - No correlation between #read
  - No correlation between #writ

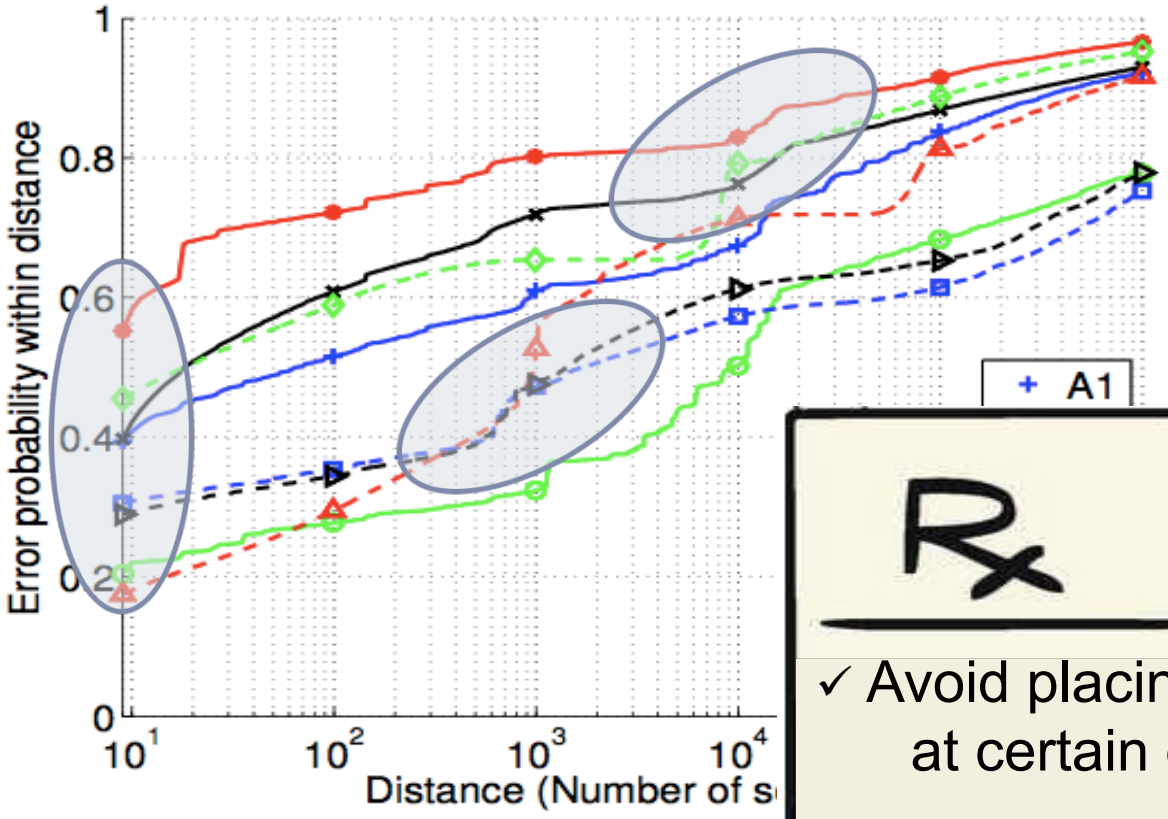


- Needs further investigation (future work).

Questions unanswered …

- What level of protection to use when?
  - What is the error probability at a given time?
- Same protection scheme across entire drive?
  - Are all parts equally likely to develop errors?
- Scrubbing potentially harmful?
  - Do additional read operations increase error rate?
- What is the common distance between errors …
  - Important for example for replica placement

How far are errors spaced apart?



- 20-60% of errors have a neighbor v
- Probability concentration (bumps)

Rx

- ✓ Avoid placing replicas at certain distances
- ✓ Explains why single parity scheme not always helpful

Questions unanswered …

- What level of protection to use when?
  - What is the error probability at a given time?
- Different protection for different parts of the drive?
  - Are all parts equally likely to develop errors?
- Scrubbing potentially harmful?
  - Do additional read operations increase error rate?
- What is the common distance between errors …
  - Important for replica placement
- Are errors that are close in space also close in time?
  - Yes!

Questions unanswered …

- What level of protection to use when?
  - What is the error probability at a given time?
- Different protection for different parts of the drive?
  - Are all parts equally likely to develop errors?
- Scrubbing potentially harmful?
  - Do additional read operations increase error rate?
- What is the common distance between errors …
  - Important for replica placement
- Are errors that are close in space also close in time?
  - Yes!
- And many other questions – see paper!

Conclusion

- Evaluated potential of different protection schemes
  - Scrubbing
    - Simple new scheme (staggered scrubbing) performs very well!
  - Intra-disk redundancy
    - Single parity can recover LSEs in 50-75% of the drives
    - Need to look at more complex schemes for coverage beyond that
- Looked at characteristics of LSEs
  - And how to exploit them for reliability
- Many characteristics not captured well by simple models
  - Provided parameters for models

- Thanks!
  - To NetApp for sharing the data
  - To you for listening
- Questions?